working paper
department
of economics

SELF-CONFIRMING EQUILIBRIUM

Drew Fudenberg
David K. Levine

No. 581 June 1991

massachusetts
institute of
technology

50 memorial drive
cambridge, mass. 02139

SELF-CONFIRMING EQUILIBRIUM

Drew Fudenberg
David K. Levine

No. 581 June 1991

SELF-CONFIRMING EQUILIBRIUM

by

Drew Fudenberg

and

David K. Levine*

June 1991

*Departments of Economics, Massachusetts Institute of Technology and University of California, Los Angeles. We thank David Kreps for many helpful conversations. National Science Foundation grants 89-07999-SES, 90-23697-SES, 88-08204-SES and 90-9008770-SES, and the Guggenheim Foundation provided financial support.

Abstract

Self-confirming equilibrium differs from Nash equilibrium in allowing players to have incorrect beliefs about how their opponents would play off of the equilibrium path. We provide several examples of ways that self-confirming and Nash equilibria differ. In games with "identified deviators," all self-confirming equilibrium outcomes can be generated by extensive-form correlated equilibria. In two-player games, self-confirming equilibria with "unitary beliefs" are Nash.

1. Introduction

Nash equilibrium and its refinements describe a situation in which (i) each player's strategy is a best response to his beliefs about the the play of his opponents, and (ii) each player's beliefs about the opponents' play are exactly correct. We propose a new equilibrium concept, self-confirming equilibrium, that weakens condition (ii) by requiring only that players' beliefs are correct along the equilibrium path of play. Thus, each player may have incorrect beliefs about how his opponents would play in contingencies that do not arise when play follows the equilibrium, and moreover the beliefs of different players may be wrong in different ways.

The concept of self-confirming equilibrium is motivated by the idea that non-cooperative equilibria should be interpreted as the outcome of a learning process, in which players revise their beliefs using their observations of previous play. Suppose that each time the game is played the players observe the actions chosen by their opponents, (or more generally, the terminal node of the extensive form) but that players do not observe the actions the opponents would have played at the information sets that were not reached along the path of play. Then, if a self-confirming equilibrium occurs repeatedly, no player ever observes play that contradicts his beliefs, so the equilibrium is "self-confirming" in the weak sense of not being inconsistent with the evidence. By analogy with the literature on the bandit problem (e.g. Rothschild [1974]) one might expect that a non-Nash self-confirming equilibrium can be the outcome of plausible learning processes. This point was made by Fudenberg and Kreps [1988], who gave an example of a learning process that converges to non-Nash outcome unless the players engage in a sufficient amount of "experimentation" with actions that do not maximize the current period's expected payoff. Our notion of self-confirming equilibrium was developed to capture the implications of learning when players do little or none of this experimentation.

To illustrate the relationship between Nash equilibrium and self-confirming equilibrium, note first that in a one-shot simultaneous-move game, every information set is reached along every path, so that self-confirming equilibrium reduces to the Nash condition that beliefs are correct at every information set. Somewhat less obvious is the fact that self-confirming equilibria must have Nash outcomes in any two-player game, so long as each player has "unitary beliefs," meaning that each strategy that the player uses with positive probability is a best response to the same (possibly incorrect) beliefs about the opponent's off-path play.

Unitary beliefs seem natural if we think of equilibrium as corresponding to the outcome of a learning model with a single player 1 and a single player 2, etc, as in Fudenberg and Kreps. We were led to consider the alternative of heterogeneous beliefs — each strategy a player uses with positive probability may be a best response to a different belief about his opponents — by our [1990], [1991] study of learning in models where a large number of individual players of each type are randomly matched with one another each period. In such models, heterogeneous beliefs can arise because different individuals have different learning experiences or different prior beliefs. When heterogeneous beliefs are allowed, the self-confirming equilibria of two-player game need not be Nash equilibria of the original game, but rather are Nash equilibria of an extended version of the game in which players can observe the outcome of certain correlating devices. Moreover, the self-confirming equilibrium outcomes are a subset of the outcomes of Forges's [1986] extensive-form correlated equilibria.

This inclusion does not obtain in general n-player games, as shown by the example of Fudenberg and Kreps. However, it does obtain in any game which has "identified deviators," meaning that deviations by different players cannot lead to the same information set. In these games, moreover, every outcome of a self-confirming equilibrium with unitary beliefs is Nash, provided that each player's subjective uncertainty corresponds to independent randomizations by his opponents. (This independence condition is difficult to explain informally; it is discussed at length on page 7.)

Since self-confirming equilibrium requires beliefs be correct along the equilibrium path of play, it is inherently an extensive-form solution concept, in contrast to Nash equilibrium, which can be defined on the strategic form of the game. Indeed, two extensive form games with the same strategic form can have different sets of self-confirming equilibria. This conclusion runs counter to the argument, recently popularized by Kohlberg and Mertens [1986], that the strategic form encodes all strategically relevant information, and two extensive forms with the same strategic form will be played in the same way. However, dependence on the extensive form is natural when equilibrium is interpreted as the result of learning, as the strategic form does not pin down how much of the opponents' strategies each player will observe when the game is played. In our view, the contrast between our approach and that of Kohlberg and Mertens shows that it is better to specify the process that leads to equilibrium play before deciding which games are equivalent or which equilibria are most reasonable.

The idea of self-confirming equilibrium with unitary beliefs is implicit in the work of Fudenberg and Kreps; our contribution here is to give a formal definition of the concept and explore its properties in various classes of games. We first noted the way that heterogeneous beliefs could allow for a form of extensive-form correlation in our [1990] paper on steady-state learning. Battagalli and Guatoli's [1988] conjectural equilibrium, and Rubinstein and Wolinsky's [1990] rationalizable conjectural equilibrium are also motivated by the idea that equilibrium corresponds to the steady state of a learning model. Their work differs from ours in considering a more general formulation of the information players observe about one another's strategies when the game is played and in restricting attention to unitary beliefs. Kalai and Lehrer's [1991] concept of a private-beliefs equilibrium assumes that beliefs are both independent and unitary; they extend our observation that such equilibria have Nash outcomes in multi-stage games by allowing for beliefs that are only approximately correct on the path of play. Canning's [1990] social equilibrium is also motivated by the idea that equilibrium corresponds to a steady state of a learning system; his concept differs in being defined in terms of the learning process, instead of being a reduced-form notion defined on the original game.

Although the motivation for self-confirming equilibrium is the idea that equilibrium is the result of learning, this paper discusses only the equilibrium concept and not its learning-theoretic foundations. Our [1991] paper considers the steady states of a system in which a fixed stage game is played repeatedly by a large population of players who are randomly matched with one another, and learn about their opponents' strategies by observing play in their own matches. Individual players remain in the population a finite number of periods; new players enter to keep the total population size constant. Entering players believe that they face a steady-state distribution on the opponents' play, and update their exogenous priors over the true steady state using Bayes rule. Given their beliefs, players choose their strategies in each period to maximize their expected present value; in particular any "experiments" must optimal even though they may have short-run costs.

Our [1991] paper shows that if lifetimes are long, then steady states approximate those of self-confirming equilibria. If in addition players are very patient, they do enough experimentation that they learn the relevant aspects of off-path play, and steady states approximate Nash equilibria.

2. The Stage Game

The stage game is an I-player extensive-form game of perfect recall. The game tree X, with nodes $x \in X$, is finite. The terminal nodes are $z \in Z \subset X$. Information sets, denoted by $h \in H$, are disjoint subsets of $X \backslash Z$. The information sets where player i has the move are $H_i \subset H$, and $H_{-i} = H \backslash H_i$ are information sets for other players. The feasible actions at information set $h_i \in H$ are denoted $A(h_i)$; $A_i = \cup_{h_i \in H_i} A(h_i)$ is the set of all feasible actions for player i, and $A_{-i} = \cup_{j \neq i} A$; are the feasible actions to player i's opponents.

A pure strategy for the player i, s_i, is a map from information sets in H_i to actions satisfying $s_i(h_i) \in A(h_i)$; S_i is the set of all such strategies. We let $s \in X = \times_{i=1}^{I} S_i$ denote a strategy profile for all players, and $s_{-i} \in S_{-i} = \times_{j \neq i} S_i$. Each player i receives a payoff in the stage game that depends on the terminal node. Player i's payoff function is denoted $u_i \colon Z \to \mathbb{R}$; each player knows his own payoff function. Let $\Delta(\cdot)$ denote the space of probability distributions over a set. Then a mixed strategy profile is $\sigma \in \times_{i=1}^{I} \Delta(S_i)$.

Let $Z(s_i)$ be the subset of terminal nodes that are reachable when s_i is played. Let $H(s_i)$ be the set of all information sets that can be reached if s_i is played.

We will also need to refer to the information sets that are reached with positive probability under σ, denoted $\bar{H}(\sigma)$. Notice that if σ_{-i} is completely mixed, then $\bar{H}(s_i,\sigma_{-i}) = H(s_i)$, as every information set that is potentially reachable given s_i has positive probability.

In addition to mixed strategies, we define behavior strategies. A behavior strategy for player i, π_i, is a map from information sets in H_i to probability distributions over moves: $\pi_i(h_i) \in \Delta(A(h_i))$, and Π_i is the set of all such strategies. As with pure strategies, $\pi \in \Pi = \times_{i=1}^{I} \Pi_i$, and $\pi_{-i} \in \Pi_{-i} = \times_{j\neq i} \Pi_j$. Let $p(z|\pi)$ be the probability that z is reached under profile Π; define $p(x|\pi)$ analogously. (Note that the probability p will reflect the probability distribution on nature's moves.)

Since the game has perfect recall, each mixed strategy σ_i induces an unique equivalent behavior strategy denoted $\pi_i(\cdot|\sigma_i)$. In other words, $\pi_i(h_i|\sigma_i)$ is the probability distribution over actions at h_i induced by σ_i.

We will suppose that all players know the structure of the extensive form, and so in particular know the strategy spaces of their opponents. We have already assumed players know their own payoff function and the probability distribution on nature's moves, so the only uncertainty each player faces concerns the strategies his opponents will play. To model this "strategic uncertainty," we let μ_i be a probability measure over Π_{-i}, the set of other players' behavior strategies. Fix s_i. Then the marginal probability of a terminal node z is

$$(2.1) \qquad p_i(z|s_i,\mu_i) = \int p_i(z|s_i,\pi_{-i})\mu_i(d\pi_{-i}) .$$

This in turn gives rise to preferences

$$(2.2) \qquad u_i(s_i,\mu_i) = \Sigma_{z\in Z(s_i)} p_i(z|s_i,\mu_i)u_i(z) \ .$$

It is important to note that even though the beliefs μ_i are over opponents' behavior strategies, and thus reflect player i's knowledge that his opponents choose their randomizations independently, the marginal distribution $p(\cdot|s_i,\mu_i)$ over terminal nodes can involve correlation between the opponents' play. For example, if players 2 and 3 simultaneously choose between U and D, player 1 might assign probability 1/4 to $\pi_2(U) = \pi_3(U) = 1$, and probability 3/4 to $\pi_2(U) = \pi_3(U) = 1/2$. Even though both profiles in the support of μ_i suppose independent randomization by players 2 and 3, the marginal distribution on their joint actions is p(U,U) = 7/16 and p(U,D) = p(D,U) = p(P,D) = 3/16, which is a correlated distribution. This correlation reflects a situation where player 1 believes some unobserved common factor has helped determine the play of both of his opponents. If, as we have supposed, the opponents are in fact randomizing independently, we should expect player 1 to learn this if he obtains sufficiently many observations. However, if few or no observations are accumulated, the correlation in the predicted marginal distribution can persist.[1]

3. Self-Confirming Equilibrium and Consistent Self-Confirming Equilibrium.

One way to define a Nash equilibrium is as a mixed profile σ such that for each $s_i \in \text{support}(\sigma_i)$ there exists beliefs μ_i such that

$$s_i \text{ maximizes } u_i(\cdot,\mu_i), \text{ and}$$

$$\mu_i(\{\pi_{-i}|\pi_j(h_j) = \pi_j(h_j|\sigma_j)\}) = 1 \quad \text{for all } h_j \in H_{-i} \ .$$

In other words, each player optimizes given his beliefs, and his beliefs are a point mass on the true distribution.

One of the goals of this paper is to introduce the notion of a self-confirming equilibrium, which weakens Nash equilibrium by relaxing the second requirement above. Instead of requiring that beliefs are correct at each information set, self-confirming equilibrium requires only that, for each s_i that is played with positive probability, beliefs are confirmed by the information revealed when s_i and σ_{-i} are played, which we take to be corresponding distribution on terminal nodes $p(s_i,\sigma_{-i})$. This corresponds to the idea that the terminal node reached is observed at the end of each play of the game.

The idea that beliefs need only be correct along the path is a natural consequence of a Bayesian approach to the formation of forecasts we study below: Bayesian learning should not be expected to lead to correct beliefs about play at information sets that are never reached.

Definition 1: Profile σ is self-confirming if for each $s_i \in \text{support}(\sigma_i)$ there exists beliefs μ_i such that

(i) s_i maximizes $u_i(\cdot,\mu_i)$, and

(ii) $\mu_i \left[\{\pi_{-i} | \pi_j(h) = \pi_j(h|\sigma_j)\}\right] = 1$ for all $s \neq i$ and $h_j \in \bar{H}(s_i,\sigma_{-i})$

Condition (ii) requires that player i's beliefs be concentrated on the subset of Π that coincide with the true distribution at information sets that are reached with positive probability when player i plays s_i. His beliefs about play at other information sets need not be concentrated on a single behavior strategy, and at these information sets his beliefs can incorporate correlation of the kind discussed in the last section. We emphasize that each $s_i \in \text{support}(\sigma_i)$ may be confirmed by a different belief

μ_i. In the definition of Nash equilibrium, this flexibility is vacuous, as each μ_i must be exactly correct; the flexibility matters once beliefs are allowed to be wrong. This diversity of beliefs is natural in a learning model with populations of each type of player: different player i's may have different beliefs, either due to different priors or to different observations.

If the same beliefs μ_i can be used to rationalize each $s_i \in \text{support}(\sigma_i)$, we will say that the equilibrium has unitary beliefs. This restriction corresponds to learning models with a single player of each type. We will occasionally speak of heterogeneous beliefs when we want to emphasize that beliefs need not be unitary. A self-confirming equilibrium is independent if for all sets $\bar{\Pi}_j \subset \Pi_j$, $\mu_i(\times_{j\neq i}\bar{\Pi}_j) = \Pi_{j\neq i}\mu_i(\bar{\Pi}_j)$ so that learning player j's behavior strategy would not change i's beliefs about $k\neq j$.

Definition 2: Profile σ is a consistent self-confirming equilibrium if for each $s_i \in \text{support}(\sigma_i)$ there are beliefs μ_i such that

(i) s_i maximizes $u_i(\cdot,\mu_i)$, and

(ii′) $\mu_i\left[\{\pi_{-i} \mid \pi_j(h_j) = \pi_j(h_j|\sigma_j)\}\right] = 1$ for all $j \neq i$ and $h_j \in H(s_i)$.

In words, self-confirming equilibrium requires that for each s_i that player i gives positive probability, player i correctly forecasts play at all information sets that will be reached with positive probability under profile (s_i,σ_{-i}). Consistent self-confirming equilibrium requires further that player i's beliefs be correct at all information sets that could possibly be reached when he plays s_i under some play of the opponents. This stronger requirement captures the information player i would obtain in a learning model if his opponents play each of their strategies sufficiently often. We call this a "consistent" equilibrium because in this case if both player i and player j can unilaterally deviate and cause information set h to be reached, then both players' beliefs about play at h are correct, and in particular are equal to each other.

Note that in a one-shot simultaneous-move game, all information sets are on the path of every profile, so the sets $H(s_i,\sigma_{-i})$ are all of H, and condition (ii) requires that beliefs be exactly correct. Hence in these games, all self-confirming equilibria are Nash. In more general games, the self-confirming equilibria can be a larger set, as shown by the examples of the next section.

4. The Characterization of Self-Confirming Equilibria

This section examines the properties of self-confirming equilibria. We begin with an example of a self-confirming equilibrium that is not consistent self-confirming. The example has the property that one player cannot distinguish between deviations by two of his opponents; we show that in the opposite case of "identified deviators" any self-confirming equilibrium is consistent self confirming. We then provide several examples of ways in which consistent self-confirming equilibria can fail to be Nash, and show that all consistent self-confirming equilibria have outcomes that can be supported by the extensive-form correlated equilibria defined by Forges [1986]. Finally, we show that consistent self-confirming equilibria with independent, unitary beliefs have the same outcomes as Nash equilibria.

Example 1 [Fudenberg-Kreps]: In the three player game illustrated in Figure 1, player 1 moves first. If he plays A, player 2 moves next; if he plays D, player 3 gets the move. If player 2 gets the move, he can either play A, which ends the game, or play D, which gives the move to player 3. The key feature of the game is that if player 3 gets the move, he cannot tell whether player 1 played D, or player 1 played A and player 2 played D.

Fudenberg and Kreps [1988] use this game to show that learning need not lead to Nash equilibrium even if players are long-lived. Suppose that player 1 expects player 3 to play R and player 2 expects player 3 to play L. Given these beliefs, it is optimal for players 1 and 2 to play A_1 and A_2. Moreover, (A_1, A_2) is a self-confirming equilibrium. However, it is not a Nash equilibrium outcome: Nash equilibrium requires players 1 and 2 to have the same (correct) beliefs about player 3's play, and if both have the same beliefs, at least one of the players must choose D. (If the beliefs assign probability more than 1/3 to L and 2 plays A, then 1 plays D, while if the beliefs assign probability more than 1/3 to R and 1 plays A then 2 plays D.)

When this example has been presented in seminars, the following question has frequently been raised: Shouldn't player 2 revise his beliefs about player 3 in the direction of 3 playing R when he sees player 1 play A? And, in the spirit of the literature on the impossibility of players "agreeing to disagree" (Aumann [1976], Geanakoplos and Polemarchakis [1982], and so forth) shouldn't players 1 and 2 end up with the same beliefs about player 3's strategy?

Our response is to note that, while this sort of indirect learning could occur in our model, it need not do. First, the indirect learning supposes that players know (or have strong beliefs about) one another's payoffs, which is consistent with our model but is not necessarily the case. Second, even if player 2 knows player 1's payoffs, and hence is able to infer that player 1 believes player 3 will play R, it is not clear that this will lead player 2 to revise his own beliefs. It is true that player 2 will revise his beliefs if he views the discrepancy between his own beliefs and player 1's as due to information that player 1 has received but player 2 has not, but player 2 might also believe that player 1 has no objective reason for his beliefs, but has simply made a mistake. The "agreeing to disagree" literature ensures all differences in beliefs are attributable to objective information by supposing that the players' beliefs are consistent with Bayesian updating from a common prior distribution. But when equilibrium is interpreted as the result of learning, the assumption of a common prior is inappropriate. Indeed, the question of whether learning leads to Nash equilibrium can be rephrased as the question of whether learning leads to common <u>posterior</u> beliefs starting from arbitrary priors. (To emphasize this point, recall that assuming players have a common prior distribution over one another's strategies is equivalent to assuming that the beliefs correspond to a correlated equilibrium (Aumann [1987]), and assuming an independent common prior is equivalent to Nash equilibrium. (Brandenburger and Dekel [1987]).

While (A_1, A_2) in Example 1 is a self-confirming equilibrium (with unitary beliefs) it is not a consistent self-confirming equilibrium, as players 1 and 2 have different beliefs about player 3's play yet player 3's information set h_3 belongs to both $H(A_1)$ and $H(A_2)$. The reason that this inconsistency matters is that both player 1 and player 2 can cause h_3 to be reached by deviating from the equilibrium path. Thus the game does not have "observed deviators" in the sense of the following definition.

<u>Definition</u> 3: A game has <u>observed</u> <u>deviators</u> if for all players i, all strategy profiles s and all deviations $s'_i \neq s_i$, $h \notin \bar{H}(s'_i, s_{-i}) \backslash \bar{H}(s)$ implies that there is no s'_{-i} with $h \in \bar{H}(s_i, s'_{-i})$.

In words, this definition says that if some deviation from s by player i leads to a new information set h, then the information set can only be reached if player i deviates. Games of perfect information satisfy this condition, as do repeated games with observed actions. More generally, the condition is satisfied by all "multi-stage games with observed actions," meaning that the extensive form can be parsed into "stages" with the properties that the beginning of each stage corresponds to a proper subgame (Selten [1975]), and that within each stage all players move simultaneously.[2] The following result shows that the condition is also satisfied in all two-player games of interest:

<u>Lemma 1</u>: Every two-player game of perfect recall has observed deviators.

<u>Proof:</u> Suppose to the contrary that there exists a profile $s = (s_1, s_2)$, and information set h such that $h \notin \bar{H}(s_1, s_2)$, but $h \in \bar{H}(s_1, s'_2)$ for some s'_2 and $h \in \bar{H}(s'_1, s_2)$ for some s'_1. If $h \in H_1$ then player 1 cannot distinguish between s_1 and s'_1, while $h \in H_2$ implies that player 2 cannot distinguish between s_2 and s'_2. ∎

<u>Theorem 1</u>: In games with observed deviators, self-confirming equilibria are

consistent self-confirming.

<u>Proof:</u> The idea of the proof is that a player's beliefs about play at information sets he cannot cause to be reached do not influence his play. Since the game has observed deviators, only one player $\hat{i}(h)$ can cause a given information set to be reached. So, starting from a self-confirming equilibrium, we construct new beliefs in which all players have the same beliefs about play at h as player $\hat{i}(h)$ did in the original equilibrium.

To make this precise, suppose that σ is a self-confirming equilibrium. Then for each player i and each $s_i \in \text{support}(\sigma_i)$ there is a μ_i that satisfies (i) and (ii) of definition 3.1. We will define new beliefs μ'_i that coincide with μ_i except on $H(s_i)\setminus \bar{H}(s_i, \sigma_{-i})$, and assign probability 1 to the true play $\pi(h|\sigma_{-i})$ at information sets in $H(s_i)\setminus\bar{H}(s_i, \sigma_{-i})$. To do this, let $Q = H(s_i)\setminus \bar{H}(s_i, \sigma_{-i})$, let Π^Q_{-i} be the projection of Π_{-i} onto Q, and let μ^Q_i be the marginal distribution μ_i induces on Π^Q_{-i}. Let Π^P_{-i} be the projection of Π_{-i} onto $\bar{H}(s_i,\sigma_{-i})$, and let μ^P_i be the distribution on Π^P_{-i} that assigns probability 1 to $\pi_j(h_j|\sigma_j)$ at each $h_j \in \bar{H}(s_i, \sigma_{-i})$. Finally, set $\mu'_i = \mu^Q_i \times \mu^P_i$.

Since μ_i satisfies condition (ii) of definition 1, μ'_i satisfies the stronger condition (ii′). Condition (ii) also implies that $\bar{H}(s_i,\mu_i) = \bar{H}(s_i,\sigma_{-i})$, that is, player i correctly predicts the equilibrium path of play when he plays s_i, and it is then clear from the definition of μ'_i that $\bar{H}(s_i, \mu'_i) = \bar{H}(s_i, \mu_i)$ as well. Moreover, information sets in $H(s_i)/ \bar{H}(s_i, \mu'_i)$ are not reached under (s_i, μ'_i), but can be reached if some player j deviates. Since the game has observed deviators, these information sets cannot be reached if none of player i's opponents deviate, that is, the information sets are not in $\bar{H}(s'_i, \mu'_i)$ for any s'_i. Hence player i's expected payoff to every action is the same under μ_i and μ'_i, so s_i is a best response to μ'_i. ■

<u>Corollary</u>: Self-confirming equilibria are consistent self-confirming in all two-player games of perfect recall.

Even consistent self-confirming equilibria need not be Nash. There are two reasons for this difference. First, consistent self-confirming equilibrium allows a player's uncertainty about his opponents' strategies to be correlated, while Nash equilibrium requires that the beliefs be a point mass on a behavior strategy profile.

<u>Example</u> <u>2</u> [Untested Correlation]: In the game in figure 2, player 1 can play A, which ends the game, or play L_1, M_1, or R_1, all of which lead to a simultaneous-move game between players 2 and 3, neither of whom observes player 1's action. In this game, A is a best response to the correlated distribution $p(L_2, L_3) = p(R_2, R_3) = 1/2$. Thus if player 1's prior beliefs are either that 2 and 3 always play L, or that they always play R, then player 1's best response is to play A, and so A is the outcome of a self-confirming equilibrium.

However, we claim that A is not a best response to any strategy profile for players 2 and 3. Verifying this is straightforward but tedious: Let p_2 and p_3 be the probabilities that players 2 and 3, respectively, assign to L_2 and L_3. In order for A to be a best response, the following 3 inequalities must be satisfied:

(4.1) $$4[p_2p_3-(1-p_2)(1-p_3)] \leq 1, \text{ or } p_2 + p_3 \leq 5/4,$$

(4.2) $4[-p_2p_3 + (1-p_2)(1-p_3)] \leq 1$, or $p_2 + p_3 \geq 3/4$, and

(4.3) $p_2(1-p_3) + (1-p_2)p_3 \leq 1/3$.

We will show that when constraints (4.1) and (4.2) are satisfied, (4.3) cannot be. For any $p_2 \leq 1/2$, the left-hand side of (4.3) is minimized when p_3 is as small as possible, that is, for $p_3(p_2) = 3/4 - p_2$. The minimized value is $2p_2^2 - 3/2\ p_2 + 3/4$, and this expression is minimized over p_2 at $p_2 = p_3 = 3/8$. At this point the left-hand side of (4.3) equals $15/36 > 1/3$. The case $p_2 > 1/2$ is symmetric.

We stress that the correlation in this example need not describe a situation in which player 1 believes that players 2 and 3 actually correlate their play. To the contrary, player 1 might be certain that they do not do so, and that she could learn which (uncorrelated) strategy profile they are using by giving them the move a single time. These competing explanations for the correlation – call them "objective" correlation and "subjective" correlation – cannot be distinguished in a static, reduced-form model of the kind considered in this paper. However, our [1991] paper on steady-state learning shows that the non-Nash outcome of example 2 can be the steady state of a learning process where players are certain that their opponents' actual play is an uncorrelated behavior profile.

In addition to untested correlation, there is another way that consistent self-confirming equilibria can fail to be Nash, which arises because the self-confirming concept allows each s_i that player i assigns positive probability to be a best response to different beliefs. This possibility allows for non-Nash play even in two-player games. The most immediate consequence of these differing beliefs is a form of convexification, as in the following example.

Example 3 [Public Randomization]: In the game in Figure 3, player 1 can end the game by moving L or he can give player 2 the move by choosing R. Player 1 should play L if he believes 2 will play D, and should play R if he believes 2 will play U. If player 1 plays R with positive probability, player 2's unique best response is to play U, so there are two Nash equilibrium outcomes, (L) and (R,U). The mixed profile ((1/2 L, 1/2 R), U) is a self-confirming equilibrium whose outcome is a convex combination of the Nash outcomes: Player 1 plays L when he expects player 2 to play D, and R when he expects 2 to play U, and when he plays L his forecast of D is not disconfirmed. (Moreover, this equilibrium is clearly independent.)

The next example shows that self-confirming equilibria in two player games can involve more than convexification over Nash equilibria. The idea is that by embedding a randomization over equilibria as in Example 3 in the second stage of a two-stage game, we can induce one player to randomize in the first stage even though such randomization cannot arise in Nash equilibrium. Moreover, this randomization may in turn cause the player's opponent to take an action that would not be a best response without it.

Example 4: The extensive form shown in Figure 3 corresponds to a two-stage game: In the first stage, players 1 and 2 play simultaneously, with player 1 choosing U or D and player 2 choosing L,M, or R. Before the second stage, these choices are revealed. In the second stage, only player 2 has a move, choosing between R ("Reward") costing both players 0, and P ("Punish") costing both players 10. The payoffs are additively separable between periods.

We claim first that in any Nash equilibrium of this game, player 1 must play a pure strategy and player 2 must play M in stage 1 with probability zero. Let q(U) be the conditional probability that player 2 plays R given that player 1 played U, and let q(D) be the probability of R conditional on D. Notice that player 1's payoffs depend only on whether 2 chooses R or P. If player 1 mixes between U and D, he must have the same expected payoff from each, so in order for player 1 to randomize, it must be that 3 + 10q(U) = 2 + 10q(D), or q(D) - q(U) = 1/10. But if both U and D have positive probability, then maximization by player 2 implies that he plays R, so q(U) = q(D) = 1, a contradiction. We conclude that player 1 must play a pure strategy, and consequently player 2 cannot play M.

Next, we consider correlated equilibrium, that is, a probability distribution over strategies with the property that for each player i and each s_i with positive probability, playing s_i is a best response to the distribution of s_{-i} conditional on s_i. If 1 plays U with probability 1, 2 must play L, while if he plays D with probability 1, 2 must play R. So in this case the probability of M is zero. On the other hand, if both U and D have positive probability and player 2 plays M with probability 1, then player 1 correctly anticipates that player 2 will respond to both U and D with M. In order for player 1 to play U in the first period, he must expect to be punished with positive probability. In other words, the outcome (U,M,P) must have positive probability. But this is impossible. If (U,M) has positive probability, player 2 cannot follow (U,M) with a positive probability of P. Thus, player 2 cannot play M with probability one, and the probability of M is bounded away from one in all correlated equilibria because the set of correlated equilibria is closed.

However, player 2 can play M with probability 1 in a self-confirming equilibrium: Let player 1's strategy be $\sigma_1 = (1/2\ U,\ 1/2\ D)$, and let player 2's strategy σ_2 be "play M in the first stage and play R in the second stage regardless of the first-period outcome." Player 2's strategy is a best response to the strategy σ_1 that player 1 is actually playing, and $U \in \text{support}(\sigma_1)$ is a best response to σ_2. The strategy $D \in \text{support}(\sigma_1)$ is not a best response to σ_2, but it is a best response to the belief that player 2 will play R if player 1 plays D and P if player 1 plays U; and when player 1 plays D his forecast of what would have happened if he had played U is not disconfirmed.

Although consistent self-confirming equilibria need not be Nash equilibria or even correlated equilibria, they are a special case of another equilibrium concept from the literature, namely the extensive-form correlated equilibria defined by Forges [1986]. These equilibria, which are only defined for games whose information sets are ordered by precedence (the usual case), are the Nash equilibria of an expanded game where an "autonomous signalling device" is added at every information set, with the joint distribution over signals independent of the actual play of the game and common knowledge to the players, and the player on move at each information set h is told the outcome of the corresponding device before he chooses his move.[3] Extensive-form correlated equilibrium includes Aumann's [1974] correlated equilibrium as the special case where the signals at information sets after stage 1 have one-point distributions and so contain no new information. The possibility of signals at later dates allows the construction of extensive-form correlated equilibria that are not correlated equilibria, as in Myerson [1986]. Another example is based on the self-confirming equilibrium we constructed in Example 4.

<u>Example 4 revisited</u>: We construct an extensive-form correlated equilibrium with the same distribution over outcomes as the self-confirming equilibrium in Example 4. The first-stage private signals describe play in that stage: There is a probability 1/2 of the signals (U,M) and (D,M) in stage 1. The strategies in stage 1 are to play the recommended action. The second-stage public signal takes on two values, U and D. The strategy for player 2 in stage 2 is to play P if player 1 played U and the second signal is D, and R otherwise. The second-stage public signal is perfectly correlated with player 1's first-stage private signal. Let us check that it is a Nash equilibrium for the players to use the strategies their signals recommend: Since player 1's signal reveals whether or not he will be punished for playing U, player 1 finds it optimal to obey his signal. Player 2's first signal is uninformative about player 1's stage 1 play, and so player 2 expects player 1 to randomize 1/2-1/2 in the first stage and thus plays M. Player 2 cannot improve on the recommended strategies in the second stage because he is only told to punish U when player 1's first signal was to play D, and if player 1 obeys his signal this will not occur. The role of the second signal is to tell player 2 when to punish player 1 without revealing player 1's play at the beginning of the first stage; if player 1's play was revealed at this point this would remove player 2's incentive to play M. Note that while the extensive-form correlated equilibrium and the self-confirming equilibrium have the same distribution over outcomes, they involve different distributions over strategies: In a self-confirming

equilibrium, if player 1 mixes between U and D, then player 2 must respond to both U and D with R; player 1 sometimes plays D because he incorrectly believes 2 will respond to U with P. In an extensive-form correlated equilibrium, each player's predictions about his opponents' strategies are on average correct, so if player 1 sometimes believes that player 2 responds to U with P then player 2 must assign positive probability to a strategy that does so.

<u>Theorem</u> <u>2</u>:

For each consistent self-confirming equilibrium of a game whose information sets are ordered by precedence, there is an equivalent extensive-form correlated equilibrium, that is, one with the same distribution over terminal nodes.

<u>Proof</u>: Let σ be consistent self-confirming, and for each $s_i \in$ support σ_i, let $\mu_i(s_i)$ be beliefs satisfying conditions (i) and (ii) of definition 1. We now expand the game by adding an initial randomizing device whose realization is partially revealed as private information at various information sets. A realization of this device is an I-vector with the i^{th} component a pair (s_i, π^i_{-i}) with $s_i \in S_i$ and $\pi^i_{-i} = (\pi_j)_{j \neq i} \in \Pi_{-i}$. The s_i follow the probability distribution σ (and in particular s_i and s_j are independent for $i \neq j$). The distribution of π^i_{-i} conditional on s is $\mu_i(s_i)$. Intuitively, profile π^i_{-i} is the way player i expects to be "punished" if he deviates from strategy s_i.

Initially each player i is told s_i. Subsequent revelations also depend upon s. At information sets on the path of s, $h \in \bar{H}(s)$, no additional information is revealed. At information sets that can be reached only if two or more players deviate from s no information is revealed. If $h_j \in \bar{H}(s'_i, s_{-i}) \backslash \bar{H}(s)$, so h_j is reached by player i's deviation, and $j \neq i$, then player j is told $\pi_j^i(h)$.

In a consistent self-confirming equilibrium, if $h_j \in \bar{H}(s'_i, s_{-i}) \backslash \bar{H}(s)$ and $h_j \in \bar{H}(s'_k, s_{-k}) \backslash \bar{H}(s)$, then $h_j \in H(s_i) \cap H(s_k)$. It follows that $\pi_j^i(h_j) = \pi_j^k(h_j)$ for $j \neq i,k$, so only one distinct signal is received by j.

Now consider the strategy profile $\hat{s}$ for the expanded game in which each player j plays s_j except at information sets (in the expanded game) where the signal $\pi_j^i(h_j)$ is received. At such information sets j plays according to $\pi_j^i(h_j)$.

By construction, $\hat{s}$ induces the same distribution over terminal nodes as σ does. If player i's opponents follow $\hat{s}$, player i will never receive an additional message, so player i is willing to play $\pi_i^j(h_i)$ at the probability-zero information sets where player j deviates and i is told $\pi_i^j(h_i)$. Moreover, given the initial message s_i, opponents' play is drawn from $\mu_i(s_i)$, and s_i is a best response to $\mu_i(s_i)$ from condition (i) in the definition of self-confirming equilibrium. Hence $\hat{s}$ is a Nash equilibrium of the expanded game. ■

Corollary: In games with identified deviators, every self-confirming equilibrium outcome is the outcome of an extensive form correlated equilibrium.

Remark: Note that not all outcomes of extensive-form correlated equilibria are the outcomes of consistent self-confirming equilibria. In particular, because self-confirming equilibria supposes that players choose their actions independently, the equilibrium path of play must be uncorrelated, so not even every correlated equilibrium outcome can be attained. This suggests that it might be possible to find an interesting and tighter characterization of consistent self-confirming equilibria; we have not been able to do so.

So far we have seen three ways in which self-confirming equilibria can fail to be Nash. First, two players can have inconsistent beliefs about the play of a third, as in example 1. Second, a player's subjective uncertainty about his opponents' play may induce a correlated distribution on their actions, even though he knows that their actual play is uncorrelated; this was the case in example 2. Finally, the fact that each player can have heterogeneous beliefs – that is, different beliefs may rationalize each $s_i \in \text{support}(\sigma_i)$ – introduces a kind of extensive-form correlation. Theorem 1 showed that in games with identified deviators, self-confirming equilibria are consistent, thus precluding the kind of non-Nash situation in example 1. The next theorem shows that the combination of off-path correlation and heterogeneous beliefs encompass all other ways that self-confirming equilibria can fail to be Nash.

<u>Theorem 3:</u> Every consistent self-confirming equilibrium with independent, unitary beliefs is equivalent to a Nash equilibrium.

<u>Proof:</u> Fix a consistent self-confirming equilibrium σ with independent, unitary beliefs. Thus for each player i, there is a μ_i such that conditions (i) and (ii′) of definition 1 are satisfied for all $s_i \in \text{support}(\mu_i)$, and μ_i is a product measure on Π_{-i}.

We will construct a new strategy profile σ' by constructing its equivalent behavior strategy profile π'. The idea is simply to change the play of all players $j \neq i$ to that given by player i's beliefs at all the information sets that can be reached if i unilaterally deviates from σ. The unitary beliefs condition implies that "player i's beliefs" are a single object; the requirement that the equilibrium is consistent ensures that this process is well-defined, as if deviations by two distinct players can lead to the same information set, then their beliefs at that information set are identical. Finally, the condition of independence says that player i's beliefs μ_i correspond to the behavior strategy profile π'_{-i}.

To explicitly define π' requires some notation.

Let $\hat{H}^i = \left[\bigcup_{j \neq i, s'_j} H(s'_j, \sigma_{-j}) \right] \backslash H(\sigma)$ be the set of information sets that can be reached if exactly one player $j \neq i$ deviates from σ, and let $\bar{H} = H \backslash (\hat{H}^i \cup \bar{H}(\sigma))$ be the information sets that can only be reached if player i or at least two other players deviate. Let $\hat{H} = \bigcup_i \hat{H}^i$ be all of the information sets that can be reached if exactly one player deviates.

For all players j, let $\pi_j'(h_j) = \pi_j(h_j|\sigma_j)$ at all $h_j \in H_j \setminus \hat{H}^j$, and let $\pi'_j(h_j) = \pi_j(h_j|\mu_i)$ at all h_j such that for some player $i \neq j$ and some s'_i, $h_j \in \bar{H}(s'_i, \sigma_{-i})$.

To verify that this construction is well-defined, we note first that if $h_j \in H_j \cap \hat{H}^j$ then there must be some player $i \neq j$ and some s'_i, such that $h_j \in \bar{H}(s'_i, \sigma_{-i})$. Thus the algorithm above specifies at least one value for π'_j at each h_j. Next we check that it assigns only one value to π'_j at each h_j. If there two players i and k and strategies s'_i, s'_k such that $h_j \in \bar{H}(s'_i, \sigma_{-i})$ and $h_j \in \bar{H}(s'_k, \sigma_{-k})$, then $h_j \in H(\sigma_i) \cap H(\sigma_k)$. Because the equilibrium is consistent and unitary, $\pi_j(h_j|\mu_i) = \pi_j(h_j|\mu_k) = \pi_j(h_j) = \pi'_j(h_j)$, so π'_j is well defined. Finally we check that profile π' is a Nash equilibrium. This verification has two steps. First, we claim that the original behavior strategy π_i is a best response to the transformed profile π'_{-i}. Note first that for all $j \neq i$, $\pi'_j(h_j) = \pi_j(h_j|\mu_i)$ at every information set that can be reached if all players but i follow profile σ; this implies that $\pi'_j(h_j) = \pi_j(h_j|\mu_i)$ at every information set that can be reached if all players but i follow π'. Next recall that player i's expected payoff to any action is unaffected by changes in his beliefs at information sets that cannot be reached if no other player deviates, and finally use the assumption that of independent beliefs to conclude that player i's payoff to each strategy s'_i under beliefs μ_i can be computed using the product of the corresponding marginal distributions $\pi_j(h_j|\mu_i)$. It then follows that player i's expected payoff to each s'_i is the same when he knows the opponents' strategies are π'_{-i} as when his beliefs are μ_i, so that π_i is a best response to π'_{-i}. To complete the proof, we note that π_i and π'_i differ only at information sets that cannot be reached unless some player $j \neq i$ deviates from π_j, so that π'_i is a best response to π'_{-i}. ■

<u>Corollary:</u> In two-player games, every self-confirming equilibrium with unitary beliefs is Nash.[4]

5. <u>Generalizations and Extensions</u>

Self-confirming equilibrium describes a situation in which players know their own payoff functions, the distribution over nature's moves, and the strategy spaces of their opponents; the only uncertainty players have is about which strategies their opponents will play. Moreover, the assumption that player's beliefs are correct along the path of play implicitly supposes that players observe the terminal node of the game at the end of each play.

These informational assumptions are what underlie our results relating self-confirming equilibrium to standard solution concepts, but in some cases, these informational assumptions are too strong. Thus it is of some interest to consider how the assumptions might be relaxed.

Battagali and Guatoli [1988] and Rubinstein and Wolinsky [1990] replace our assumption that players observe the terminal node of the game with a more general formulation of what the players observe when the game is played. In our view these observations should not be more informative than the terminal node of the game, and should at least allow each player to compute his own payoff. (In these models, where each player i observes signal $g_i(s)$ when the profile s is played, this constraint would require player i's utility function u_i to be measurable with respect to g_i.) It would be interesting to see a characterization of self-confirming equilibrium for the case in which each player's end-of-stage information is precisely his own payoff; the key would be finding a tractable description of how much information the payoffs convey. Another interesting case is that of games of incomplete information, with the assumption that each player observes the entire sequence of play and his own type, but not the types of his opponents. We conjecture that if each player's payoff function does not depend on his opponents' types, the set of self-confirming equilibria is the same whether or not the opponents' types are observed at the end of each round.

The other informational assumptions of self-confirming equilibrium can be relaxed as well. It is easy to generalize self-confirming equilibrium to allow for players to not know the distribution of nature's moves; see our [1990] working paper for the details. Allowing for the possibility that players do not know the extensive-form structure of the game is more difficult. One issue is that, when the extensive form is unknown, players may believe that some opponents can condition their play information the opponents cannot in fact possess. Also, in some formulations of the players' inference processes, players may become convinced that their opponents' play is influenced by the actions the player means to take at information sets that are in fact not reached. Fudenberg and Kreps [1991] discuss some of these problems, but are unable to provide a satisfactory resolution of them.

Footnotes

[1]We thank Robert Aumann for convincing us of the importance of this kind of subjective correlation.

[2]See Fudenberg and Tirole [1991] for a more detailed explanation of multi-stage games; we introduced the definition in Fudenberg and Levine [1983]. Note that the extensive form in example 2 below is not a multi-stage game with observed actions, but is a game with observed deviators. Moreover, splitting player 1's information into two consecutive choices, the first one being A or ~A, yields a multi-stage game with observed actions that has the same reduced normal form and the same set of self-confirming equilibria. This emphasizes that from the viewpoint of self confirming equilibria, identified deviators is the more fundamental property.

[3]Forges shows, in the spirit of the revelation principle, that it suffices to work with a smaller set of signalling devices. She also defines "communications equilibria," which allow the players to send messages in the course of play that influence subsequent signals.

[4]This is proved directly in Fudenberg and Kreps [1991].

References

Aumann, R. [1974], "Subjectivity and Correlation in Randomized Strategies," Journal of Mathematical Economics, 1, 67-96.

---------- [1976], "Agreeing to Disagree," Annals of Statistics, 4: 1236-1239.

---------- [1987], "Correlated Equilibrium as an Expression of Bayesian Rationality," Econometrica, 55, 1-18.

Battagalli, P. and P. Guatoli [1988], "Conjectural Equilibrium," mimeo.

Brandenburger, A. and E. Dekel [1987], "Rationalizability and Correlated Equilibria," Econometrica, 55: 1391-1402.

Canning, D. [1990], "Social Equilibrium," mimeo, Cambridge University.

Forges, F. [1986], "An Approach to Communications Equilibrium," Econometrica, 54, 1375-1386.

Fudenberg, D., and D. Kreps [1988], "A Theory of Learning, Experimentation and Equilibrium in Games," mimeo, Stanford.

---------- [1991], "A Theory of Learning, Experimentation and Equilibrium in Games," mimeo.

Fudenberg, D., and D. Levine [1983], "Subgame-Perfect Equilibria of Finite and Infinite Horizon Games," Journal of Economic Theory, 31, 251-268.

---------- [1990], "Steady State Learning and Self-Confirming Equilibrium," mimeo.

---------- [1991], "Steady State Learning and Nash Equilibrium," mimeo.

Geanakoplos, J. and H. Polemarchakis [1982], "We Can't Disagree Forever," Journal of Economic Theory, 28: 192-200.

Kalai, E. and E. Lehrer [1991], "Private Beliefs Equilibrium," mimeo.

Kohlberg, E., and J.F. Mertens [1986], "On the Strategic Stability of Equilibrium," Econometrica, 54, 1003-1034.

Myerson, R.B. [1986], "Multistage Games With Communication," Econometrica, 54, 323-358.

Rothschild, M. [1974], "A Two-Armed Bandit Theory of Market Pricing," Journal of Economic Theory 9: 195-202.

Rubinstein, A. and A. Wolinsky [1990], "Rationalizable Conjectural Equilibrium Between Nash and Rationalizability," mimeo.

Selten, R. [1975], "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," International Journal of Game Theory, 4, 25-55.

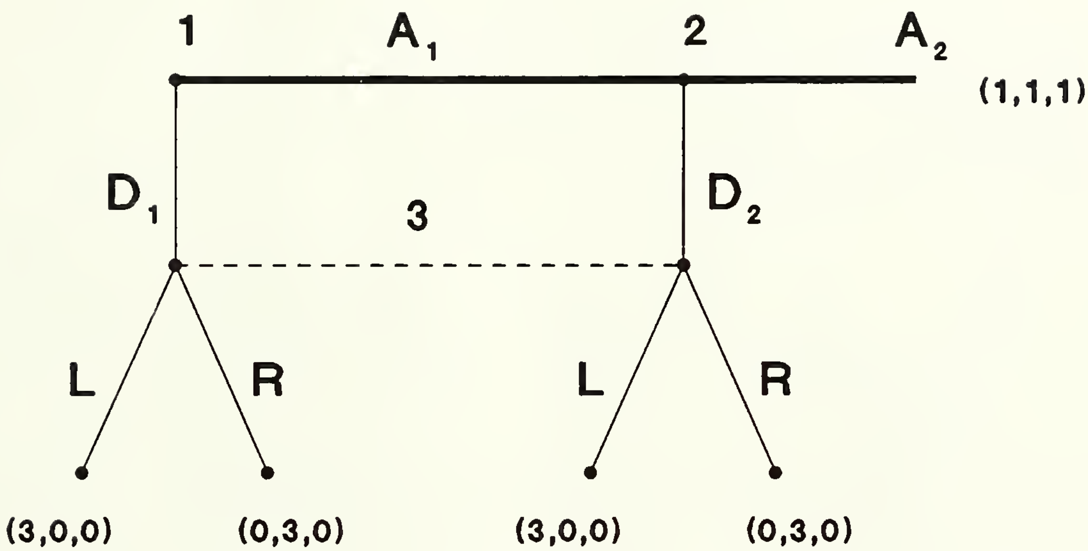
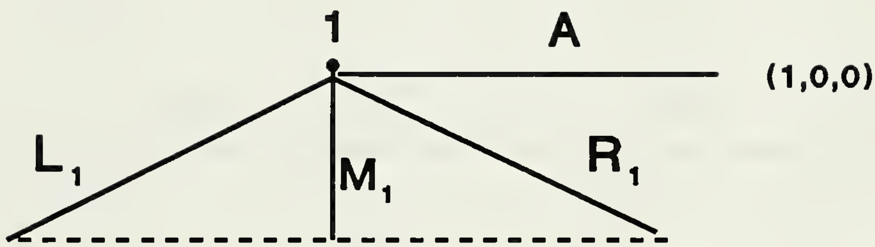


Figure 1



$$
\begin{array}{c c}
 & \begin{array}{cc} L_3 & R_3 \end{array} \\
\begin{array}{c} L_2 \\ R_2 \end{array} & \begin{bmatrix} 4,1,-1 & 0,-1,1 \\ 0,-1,1 & -4,1,-1 \end{bmatrix}
\end{array}
\quad
\begin{array}{c}
\begin{array}{cc} L_3 & R_3 \end{array} \\
\begin{bmatrix} -4,1,-1 & 0,-1,1 \\ 0,-1,1 & 4,1,-1 \end{bmatrix}
\end{array}
\quad
\begin{array}{c}
\begin{array}{cc} L_3 & R_3 \end{array} \\
\begin{bmatrix} 0,1,-1 & 3,-1,1 \\ 3,-1,1 & 0,1,-1 \end{bmatrix}
\end{array}
$$

Figure 2

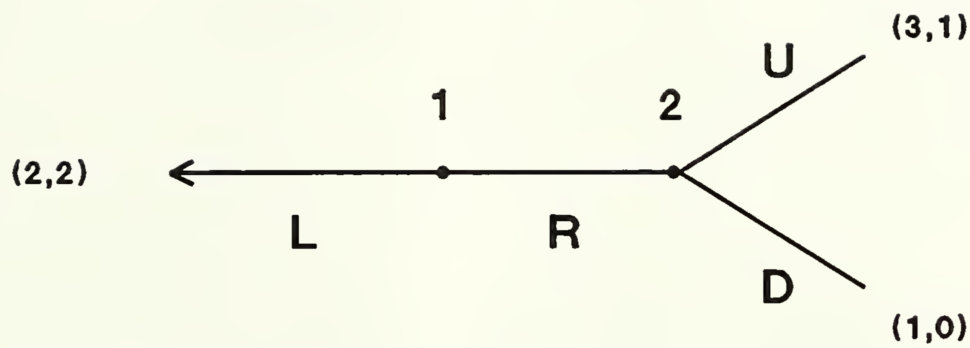


Figure 3

	L	M	R
U	13,15	13,14	13,11
D	12,11	12,14	12,15

Stage 1

R	P
0,0	-10,-10

Stage 2

Figure 4